DØ note 4874-CONF

Top Quark Mass Measurement with the Matrix Element Method in the Lepton+Jets Final State at DØ Run II

The DØ Collaboration
URL http://www-d0.fnal.gov
(Dated: July 19, 2005)

We present a measurement of the top quark mass with the matrix element method in the lepton+jets final state. As the jet energy scale represents the dominant source of systematic uncertainty, the matrix element likelihood is extended by an additional "JES" parameter, which is defined as a global scale factor relative to the Monte Carlo reference scale. The top mass is obtained from a correlated two-dimensional fit, which yields both the statistical and the systematic jet energy scale uncertainty. Using a data set of $320\,\mathrm{pb}^{-1}$ of DØ Run II data, the mass of the top quark is measured to be

$$m_{\mathrm{top}}^{\ell+\mathrm{jets}} \;=\; 169.5 \pm 4.4\,(\mathrm{stat.}+\mathrm{JES})\,^{+1.7}_{-1.6}\,(\mathrm{syst.})\,\mathrm{GeV}/c^2\ .$$

The measurement yields $JES = 1.034 \pm 0.034$, indicating good consistency between the jet energy scales in data and simulation.

Preliminary Results for the EPS Conference, Summer 2005

I. INTRODUCTION

The top quark mass measurement with the matrix element method at DØ in Run I [1] improved the statistical and systematic uncertainty significantly with respect to previous measurements and generated considerable excitement. The method underwent significant improvements for the Run II measurement presented here, most notably the addition of a second parameter to the likelihood function which addresses the jet energy scale as the dominant source of systematic uncertainty.

In $p\bar{p}$ collisions at $\sqrt{s} = 1.96\,\text{TeV}$, top quarks are predicted in the Standard Model to be produced dominantly as top-antitop pairs via $q\bar{q}$ annihilation (85 %) and gluon fusion (15 %). Both top and antitop are predicted to decay almost exclusively to a W boson and a b quark. If one of the W bosons decays hadronically to a pair of light quarks, while the other decays to either an electron or muon and the corresponding neutrino, the event is referred to as a *lepton+jets* (ℓ+jets) event. The signature of this decay in the detector is the presence of four or more jets, an isolated lepton, and missing transverse energy $\not\!\!E_{\text{T}}$ from the undetected neutrino. The dominant physics background to this process is the electroweak production of a leptonically decaying W in association with four or more quarks and gluons. Additional instrumental background arises from multi-jet events, where either a heavy flavor jet decays semi-leptonically but only the muon is reconstructed (μ+jets channel) or a jet is misidentified as an electron ($e+$jets channel). This instrumental background is referred to as "QCD" background throughout this note and is expected to be small and of similar topology as $W(\rightarrow l\nu) + \text{jets}$.

II. THE DØ DETECTOR

The DØ detector has a central-tracking system, consisting of a silicon microstrip tracker (SMT) and a central fiber tracker (CFT), both located within a 2 T superconducting solenoidal magnet [2], with designs optimized for tracking and vertexing at pseudorapidities $|\eta| < 3$ and $|\eta| < 1.6$, respectively. A liquid-argon and uranium calorimeter has a central section (CC) covering pseudorapidities $|\eta|$ up to ≈ 1.1, and two end calorimeters (EC) that extend coverage to $|\eta| \approx 4.2$, with all three housed in separate cryostats [3]. An outer muon system, at $|\eta| < 2$, consists of a layer of tracking detectors and scintillation trigger counters in front of 1.8 T toroids, followed by two similar layers after the toroids [4].

Trigger and data acquisition systems are designed to accommodate the luminosities of Run II. Based on preliminary information from tracking, calorimetry, and muon systems, the output of the first level of the trigger is used to limit the rate for accepted events to $\approx 2\,\text{kHz}$. At the next trigger stage, with more refined information, the rate is reduced further to $\approx 1\,\text{kHz}$. The third and final level of the trigger, with access to all the event information, reduces the output rate to $\approx 50\,\text{Hz}$, which is written to tape.

III. EVENT SELECTION

The event selection is mainly adopted from the topological top quark cross section analyses in the $e+$jets and $\mu+$jets channels at DØ for Run II [5]. Events are selected requiring an isolated energetic charged lepton ($p_{\text{T}} > 20\,\text{GeV}/c$, electron: $|\eta| < 1.1$; muon: $|\eta| < 2$), significant missing transverse energy ($\not\!\!E_{\text{T}} > 20\,\text{GeV}$), and exactly four good calorimeter jets ($p_{\text{T}} > 20\,\text{GeV}/c$, $|\eta| < 2.5$). A $\Delta\phi$ cut between the lepton and $\not\!\!E_{\text{T}}$ is imposed to reject events in which the transverse energy imbalance originates from a poor lepton energy measurement. The official certified DØ jet energy corrections are applied to the jets in the event. The calorimeter response for these corrections is measured in $\gamma+$jets data. The measurement is performed for the central and forward calorimeters individually. Once the scale is derived, another iteration resolves more subtle features of the jet energy correction as a function of η. The scale is applied to the jets in a $\gamma+1$ jet sample, and the variable

$$\Delta S = \frac{p_{\text{T}}^{jet} - p_{\text{T}}^{\gamma}}{p_{\text{T}}^{\gamma}} \tag{1}$$

reveals additional structure of the scale as a function of pseudorapidity. These η-dependent corrections are applied to the jets and propagated to $\not\!\!E_{\text{T}}$ before the selection of the sample.

This selection yields 70 and 80 events for the $320\,\text{pb}^{-1}$ DØ Run II data sample in the $e+$jets and $\mu+$jets channel respectively. Application of a topological likelihood technique, which was developed at DØ in Run I and refined in Run II to measure the top quark cross section [5], yields estimates on the signal fraction as well as the contribution from instrumental background (QCD) in the sample. The fits are shown in Figure 1, and the results are summarized in Table I.

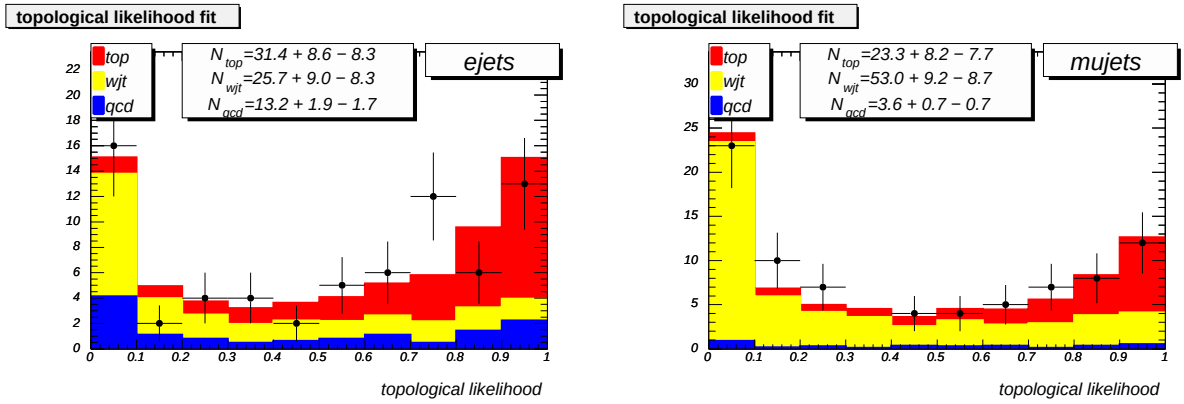

FIG. 1: Topological likelihood fit applied to the 320 pb^{-1} DØ Run II data sample. The technique is taken from the DØ Run II cross section measurements, and the event selection of this analysis is taken into account.

channel	N_{evts}	f_{top}	N_{top}	f_{QCD}
$e+$jets	70	$44.9\ ^{+12.3}_{-11.9}$ %	31.4	$18.4\ ^{+2.7}_{-2.5}$ %
μ+jets	80	$29.1\ ^{+10.3}_{-\ 9.6}$ %	23.3	$4.5\ ^{+0.9}_{-0.8}$ %
ℓ+jets	150	$36.4\ ^{+11.3}_{-10.7}$ %	54.7	$11.2\ ^{+1.8}_{-1.6}$ %

TABLE I: Composition of the $e+$jets, μ+jets, and ℓ+jets data samples, estimated with a topological likelihood technique.

IV. THE MATRIX ELEMENT METHOD

The probability of any event to be produced via a certain process is proportional to the differential cross section of that process. Consequently, the probabilities P_{sgn} and P_{bkg} for events to originate from the $t\bar{t}$ signal and the $W+$jets background process are calculated based on the respective matrix elements $\mathscr{M}_{t\bar{t}}$ and $\mathscr{M}_{W+\mathrm{jets}}$. The leading order matrix element is taken for $q\bar{q} \to t\bar{t}$ production, and $W+$jets events are described with the VECBOS [9] parameterization of the matrix element. The energy resolutions for muons and jets are taken into account as well: Transfer functions $W(E_j, E_p)$ are derived from Monte Carlo, which describe the probability for a parton with energy E_p to be reconstructed with E_j in the detector. Jet and lepton angles and electron energies are assumed to be well measured, and the probabilities are obtained by integrating over all possible parton states, where each state is weighted by its probability to produce the observed measurement. The integrations are performed with the Monte Carlo integration algorithm VEGAS [6, 7] as provided by the GNU Scientific Library [8]. All relevant jet permutations and neutrino solutions are considered, and the integration time for one jet permutation varies from 1 to 6 seconds. The total event probability is defined by combining both probabilities according to

$$P_{\mathrm{evt}}(x; m_{\mathrm{top}}, JES) = f_{\mathrm{top}} \cdot P_{\mathrm{sgn}}(x; m_{\mathrm{top}}, JES) + (1 - f_{\mathrm{top}}) \cdot P_{\mathrm{bkg}}(x; JES)\ , \tag{2}$$

where x denotes all kinematic variables of the reconstructed lepton and jets. The transverse momentum of the neutrino is obtained from the p_T imbalance of the five detected final state objects. f_{top} is the signal fraction in the sample under study. The signal probability is sensitive to the jet energy scale parameter JES, because the mass of the hadronically decaying W boson is constrained in the $t\bar{t}$ matrix element.

In order to extract the top quark mass from a set of n events with measurements $x_1, .., x_n$, a likelihood function is built from the event probabilities,

$$L(x_1, .., x_n; m_{\mathrm{top}}, JES) = \prod_{i=1}^{n} P_{\mathrm{evt}}(x_i; m_{\mathrm{top}}, JES)\ , \tag{3}$$

and evaluated for different hypotheses of m_{top} and JES. The top quark mass is determined by minimizing

$$-\ln L(x_1, .., x_n; m_{\mathrm{top}}, JES) = -\sum_{i=1}^{n} \ln(P_{\mathrm{evt}}(x_i; m_{\mathrm{top}}, JES)) \tag{4}$$

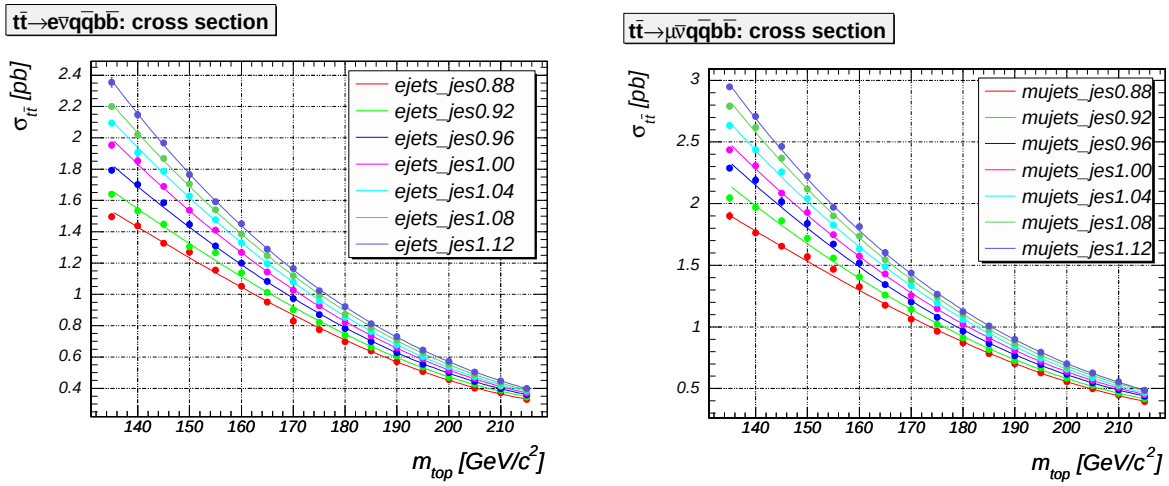

FIG. 2: Normalization of $P_{\rm sgn}$ for $e+$jets (left) and μ+jets (right) events as a function of the top quark mass $m_{\rm top}$ for different choices of the JES scale factor.

with respect to $m_{\rm top}$ and JES simultaneously, taking all correlations between both parameters into account. The signal fraction $f_{\rm top}$ is fitted simultaneously as well.

The signal probability is normalized by computing the integral of $P_{\rm sgn}$ over the 16-dimensional parton phase space as a function of $m_{\rm top}$ and JES. The kinematic selection is taken properly into account by additional integrations over the reconstructed object energies, where the transfer functions are used to relate detector cuts to parton phase space. The results are shown in Figure 2 for $e+$ jets and μ+jets events as a function of $m_{\rm top}$ for various choices of the JES scale factor.

The background probability $P_{\rm bkg}$ is calibrated such that the fitted signal fraction $f_{\rm top}$ yields the true $t\bar{t}$ fraction in Monte Carlo ensemble tests on average for all top quark mass samples.

V. CALIBRATION OF THE METHOD

Monte Carlo events which have been run through the full simulation of the DØ detector are used to derive the calibration of the mass fit. Here, $t\bar{t}$ samples with top quark pole masses of 160, 170, 175, 180, and 190 GeV/c^2 and a $W+$jets sample are used. In addition, samples with $m_{\rm top} = 175\,{\rm GeV}/c^2$ and all jets scaled by 0.92, 0.96, 1.04, and 1.08 are prepared in order to calibrate the JES fit. For each sample and each lepton channel ($e+$jets and μ+jets), $P_{\rm sgn}$ and $P_{\rm bkg}$ are calculated for 1000 events which pass the kinematic selection. 1000 ensembles are drawn from these event pools, where each event is allowed to appear in more than one ensemble and more than once in the same ensemble. This technique allows for the derivation of the expected uncertainty and the pull with better precision [10]. The ensembles are composed according to the estimates yielded by the topological likelihood fit, summarized in Table I. The QCD contribution however is substituted by $W+$jets events. Contamination of the sample with multijet events is treated as a systematic uncertainty.

The calibration derived for the ℓ+jets sample is shown in Figure 3. The pull width for both $m_{\rm top}$ and JES is in good agreement with 1.0, indicating a trustworthy error estimate by the likelihood procedure. Figure 4 illustrates that the fitted top mass does not depend on the true jet energy scale in the sample. The calibrations of both parameters are applied to the result obtained from the data sample in the following section.

VI. RESULT

The matrix element method is applied to the $320\,{\rm pb}^{-1}$ ℓ+jets data set collected at DØ during Run II. The calibrations for $m_{\rm top}$ derived in the previous section are taken into account. Although pull width deviations from 1.0 are not significant, the statistical uncertainty yielded by the mass fit is inflated accordingly. The calibrated fit result for the combined ℓ+jets sample is shown in Figure 5. The top mass is measured to be

$$m_{\rm top}^{\ell+{\rm jets}} \;=\; 169.5 \pm 4.4\,({\rm stat.} + {\rm JES})\,{\rm GeV}/c^2\ .$$

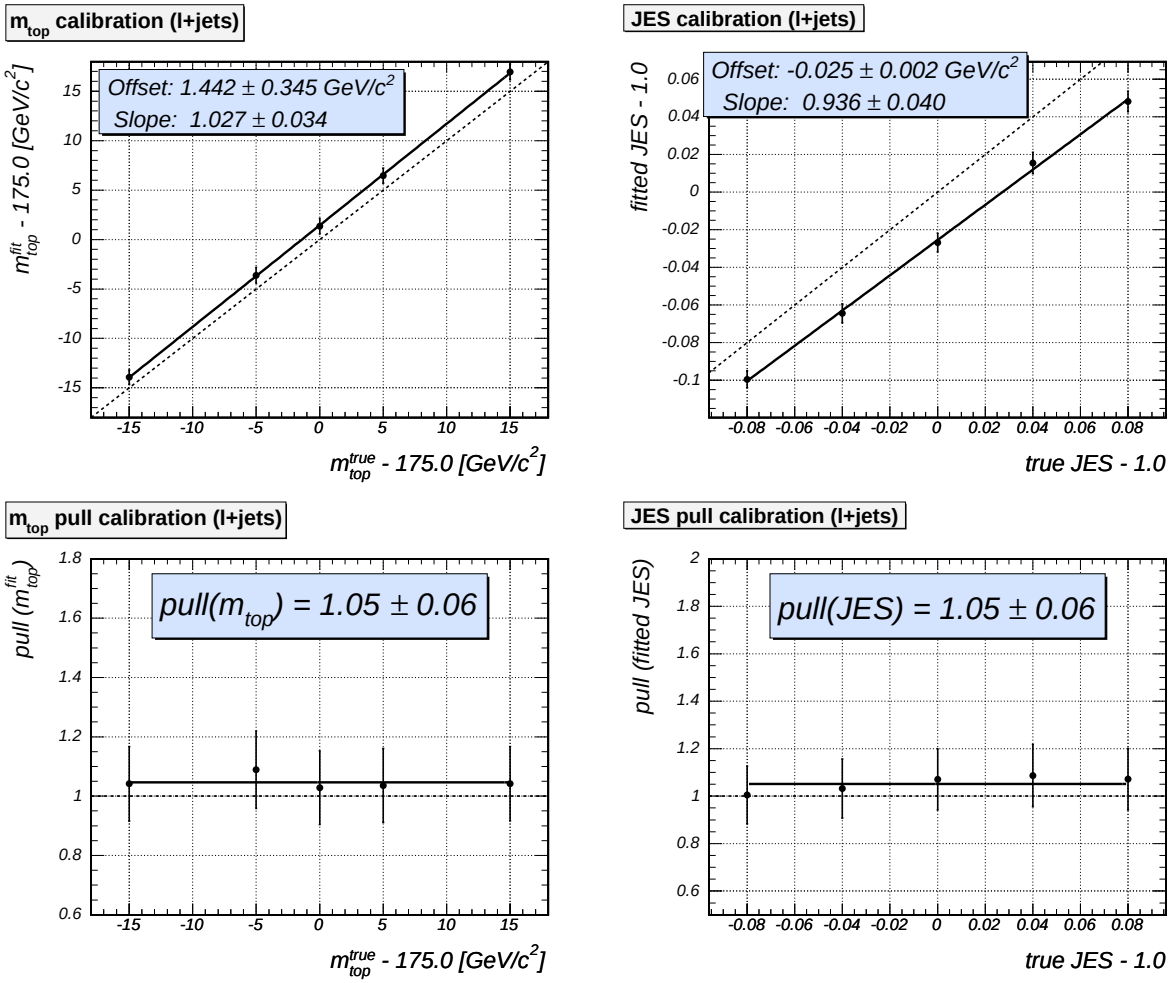

FIG. 3: Calibration of the matrix element mass fitting procedure for ℓ+jets events.

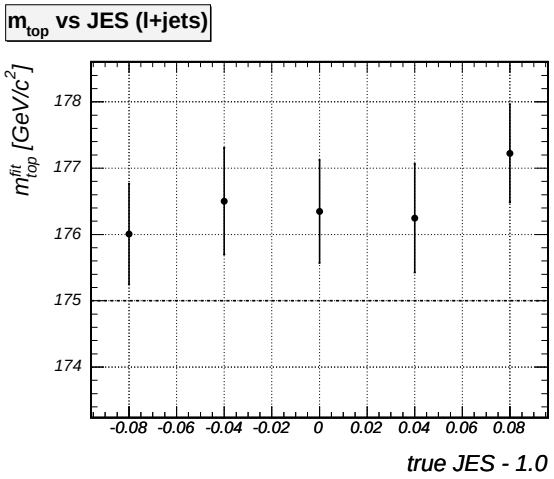

FIG. 4: Fitted $m_{\rm top}$ as a function of true jet energy scale JES: the mass fit is stable.

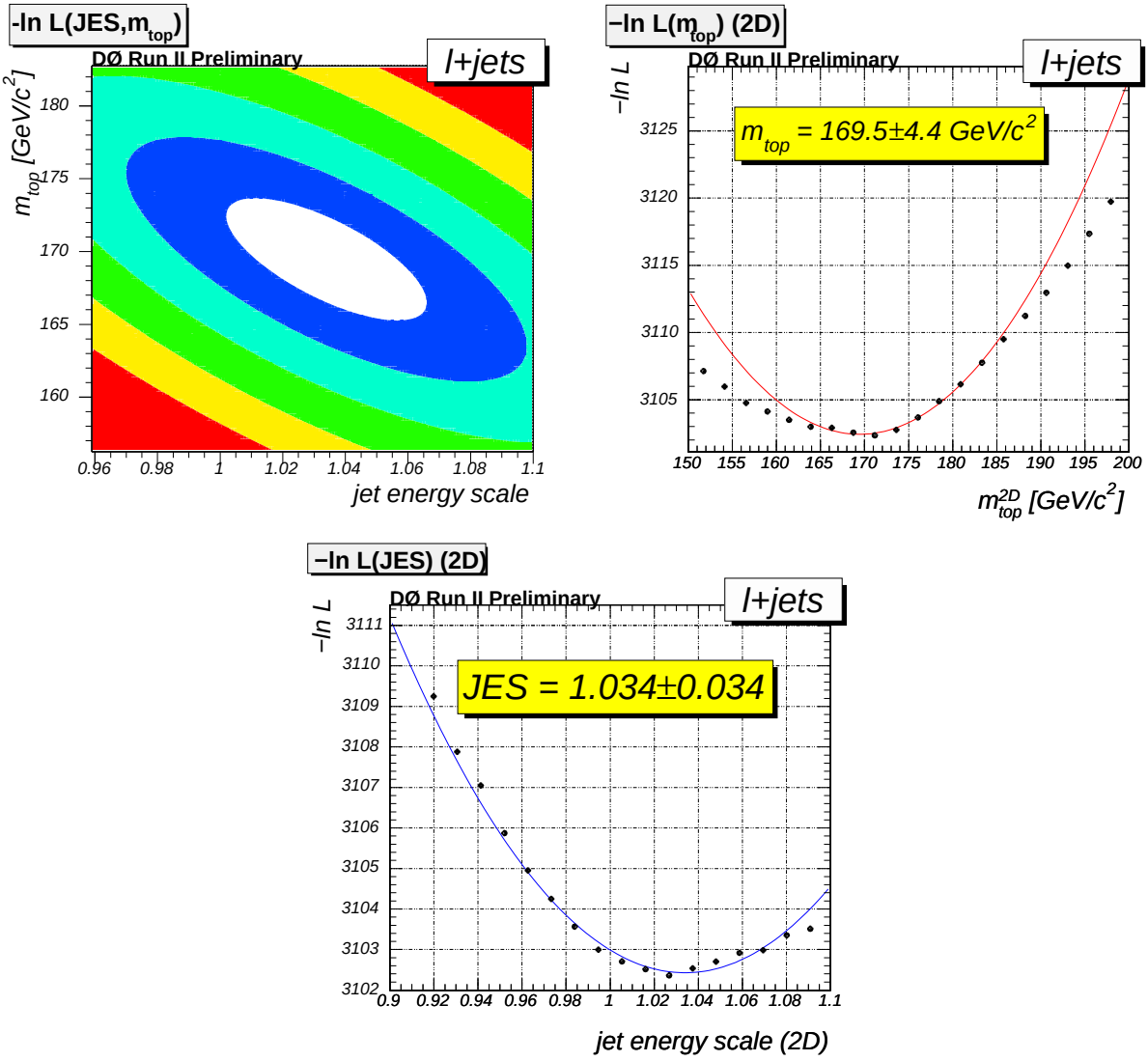

FIG. 5: Application of the matrix element method to the 320 pb^{-1} ℓ+jets data set. The $m_{\rm top}$ and JES axes correspond to the calibrated values. The top left plot shows the two-dimensional $m_{\rm top}$-JES fit with σ contours. The projection to the $m_{\rm top}$ parameter is shown in the top right plot, taking correlations into account. The corresponding $-\ln L$ points are shown as well. Similarly, the projection to the JES parameter is shown in the plot below together with the corresponding $-\ln L$ points, again taking correlations into account.

For a fixed jet energy scale, the statistical error of the fit is 3.0 GeV/c^2; thus the component from the jet energy scale uncertainty is 3.2 GeV/c^2.

The fit yields a signal fraction $f_{\rm top}$ of $0.316^{+0.049}_{-0.055}$, in good agreement with the expectation. The fitted jet energy scale of 1.034 ± 0.034 indicates that the scale in the simulation is consistent with that in the data.

As a cross-check, the data events are also considered with their jet energies scaled by a factor of 1/1.034 according to the above fit result. The missing transverse energy is adjusted in each event. Of these scaled events, 132 pass the event selection and are used to determine the $-$ln(likelihood) as a function of the W and top masses. The best $-$ln(likelihood) value as a function of the W mass is shown in Figure 6. The fitted W mass of $m_W = 80.1^{+5.1}_{-3.4}$ GeV/c^2 is in good agreement with the expectation of the value $m_W = 80.4$ GeV/c^2 assumed in the matrix element.

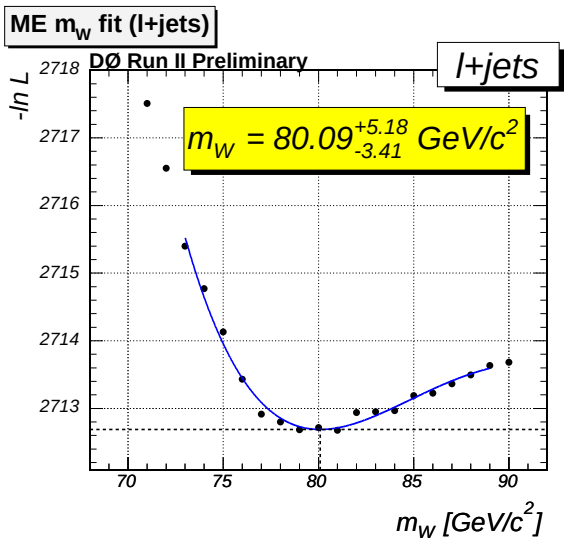

FIG. 6: Fit to the $-\ln$(likelihood) as a function of m_W for the sample of 132 events with jet energies scaled by 1/1.034.

Uncertainty	ℓ+jets [GeV/c^2]
JES p_T dependence	± 0.7
b fragmentation	± 0.71
b response (h/e)	$+0.87 - 0.75$
signal modeling	± 0.34
background modeling	± 0.32
signal fraction	$+0.50 \ -0.17$
QCD contamination	± 0.67
MC calibration	± 0.38
trigger	± 0.08
PDF uncertainty	± 0.07
TOTAL	$+1.7 \ -1.6$

TABLE II: Summary of systematic uncertainties.

VII. SYSTEMATIC UNCERTAINTIES

The jet energy scale uncertainty is already included in the error yielded by the likelihood. By fixing the JES hypothesis to 1.0 and extracting a mass value without the correlation, the contribution from jet energy scale is estimated to be $3.2\,\mathrm{GeV}/c^2$. Table II summarizes all other systematic uncertainties on the top mass measurement with the matrix element method. The total systematic uncertainty on the top mass measurement is obtained by adding all contributions in quadrature:

$$(\Delta m_{\mathrm{top}})^{\ell+\mathrm{jets}}_{syst} \ = \ +1.7 \ -1.6\,\mathrm{GeV}/c^2 \ . \tag{5}$$

Note that no systematic uncertainty is quoted due to multiple interactions/uranium noise as opposed to the Run I measurement. The effect is much smaller in Run II as a consequence of the reduced integration time in the calorimeter readout. It is moreover covered by the jet energy scale uncertainty, as the offset correction is computed seperately for data and Monte Carlo in Run II, accounting for effects arising from electronic noise and pileup.

- **JES p_{T} dependence:** The relative difference between the jet energy scales in data and Monte Carlo is fitted with a global scale factor, and the corresponding uncertainty is included in the quoted (stat. + JES) error. Any discrepancy between data and simulation other than a global scale difference may lead to an additional uncertainty on the top quark mass. This uncertainty is currently estimated to be $\pm 0.7\,\mathrm{GeV}/c^2$.

- **relative b/light jet energy scale:** While the overall jet energy scale uncertainty is included in the statistical uncertainty from the fit, differences in the b/light jet energy scale ratio between data and simulation may still

affect the measurement. Possible effects from such differences are studied using simulated $t\bar{t}$ events with different fragmentation models for b jets [13, 14]. The resulting uncertainty is $\pm 0.71\,\mathrm{GeV}/c^2$. In addition, variations of the h/e calorimeter response are considered, leading to an uncertainty of $+0.87 - 0.75\,\mathrm{GeV}/c^2$.

- **signal modeling:** When $t\bar{t}$ events are produced in association with a jet, the additional jet can be misinterpreted as a product of the $t\bar{t}$ decay. The $t\bar{t}$ system may then have significant transverse momentum, in contrast to the assumption made in the calculation of P_{sgn}. In spite of the event selection that requires exactly four jets, these events can be selected if one of the jets from the $t\bar{t}$ decay is not reconstructed. To assess the uncertainty from modeling of such events, 30% of simulated signal events (according to the difference between cross-section calculations in leading and next-to-leading order) have been replaced by events from a dedicated $t\bar{t}$+jets simulation. The observed difference between results obtained with the two signal models is symmetrized and assigned as a systematic uncertainty. The uncertainty assigned is ± 0.34 GeV/c^2.

  In addition, simulated $gg \to t\bar{t}$ and $q\bar{q} \to t\bar{t}$ events have been compared. No significant difference between the top mass calibration curves for the two processes has been found, and thus no additional uncertainty on the result is assigned.

- **background modeling:** In order to study the sensitivity of the measurement to the choice of background model, the standard W + jets Monte Carlo sample is replaced by an alternative sample with the default factorization scale of $Q^2 = m_W^2 + \sum_j p_{T,j}^2$ replaced by $Q'^2 = \langle p_{T,j} \rangle^2$. The difference obtained between ensembles constructed with the two simulations is ± 0.32 GeV/c^2 and is assigned as a systematic uncertainty.

- **signal fraction:** The signal fraction f_{top} is slightly overestimated for low true signal fractions, which leads to a small bias in the resulting top mass. The signal fraction in the data sample is varied within the uncertainties determined from the topological likelihood fit, and the resulting variation of the top mass of $+0.50\ -0.17$ GeV/c^2 is taken as systematic uncertainty.

- **QCD background:** The W + jets simulation is used to model the small QCD background in the selected event sample in the analysis. The systematic uncertainty from this assumption is computed by selecting a dedicated QCD-enriched sample of events from data by inverting the lepton isolation cut in the event selection. The calibration of the method is repeated with ensembles formed where these events are used to model the QCD background. The difference of ± 0.67 GeV/c^2 to the default calibration is assigned as a systematic uncertainty.

- **MC calibration:** The statistical uncertainty on the calibration curves shown in Figure 3 is propagated through the analysis and yields a systematic uncertainty of ± 0.38 GeV/c^2.

- **trigger:** The trigger efficiencies used in the ensemble testing procedure are varied by their uncertainties and the uncertainties from all variations are summed in quadrature. The associated systematic uncertainty is ± 0.08 GeV/c^2.

- **PDF uncertainty:** Leading-order matrix elements are used to calculate both P_{sgn} and P_{bkg}. Consequently, both calculations evaluate a leading order parton distribution function (PDF): CTEQ5L [15]. To study the systematic uncertainty on m_{top} due to this choice, the variations provided with the next-to-leading-order PDF set CTEQ6M [16] are used (no variations for a leading order PDF are available), and the result obtained with each of these variations is compared with the result using the default CTEQ6M parametrization. All differences are added in quadrature and yield a systematic uncertainty of ± 0.07 GeV/c^2.

VIII. CONCLUSION

Applying the matrix element method to a $320\,\mathrm{pb}^{-1}$ data set recorded with the DØ experiment at the Run II Tevatron, we measure the top quark mass in lepton+jets $t\bar{t}$ events to be

$$m_{\mathrm{top}}^{\ell+\mathrm{jets}} \;=\; 169.5 \pm 4.4\,(\mathrm{stat.} + \mathrm{JES})\,^{+1.7}_{-1.6}\,(\mathrm{syst.})\,\mathrm{GeV}/c^2\ .$$

The uncertainty on the jet energy scale is included in the statistical error, as the overall jet energy scale relative to the simulation is determined simultaneously with the top quark mass. We find $JES = 1.034 \pm 0.034$, indicating good consistency with the simulation.

Acknowledgments

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); FASI, Rosatom and RFBR (Russia); CAPES, CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); KRF (Korea); CONICET and UBACyT (Argentina); FOM (The Netherlands); PPARC (United Kingdom); MSMT (Czech Republic); CRC Program, CFI, NSERC and WestGrid Project (Canada); BMBF and DFG (Germany); SFI (Ireland); Research Corporation, Alexander von Humboldt Foundation, and the Marie Curie Program.

[1] V.M. Abazov *et al.*, *A precision measurement of the mass of the Top quark*, Nature 429:638-642 (2004).
[2] DØ Collaboration, V. Abazov *et al.*, *The Upgraded DØ Detector*, in preparation for submission to Nucl. Instrum. Methods Phys. Res. A.
[3] DØ Collaboration, S. Abachi *et al.*, Nucl. Instrum. Methods Phys. Res. A 338, 185 (1994).
[4] V.M. Abazov *et al.*, *The Muon System of the Run II DØ Detector*, physics/0503151 (2005).
[5] V.M. Abazov *et al.* [DØ Collaboration], *Measurement of the ttbar Production Cross Section in ppbar Collisions at sqrt(s) = 1.96 TeV using Kinematic Characteristics of Lepton+Jets Events*, hep-ex/0504043 (2005).
[6] G.P. Lepage, *A New Algorithm for Adaptive Multidimensional Integration*, Journal of Computational Physics 27:192-203, (1978).
[7] G.P. Lepage, VEGAS*: An Adaptive Multi-dimensional Integration Program*, Cornell preprint CLNS:80-447, (1980).
[8] M. Galassi *et al.*, *GNU Scientific Library Reference Manual (2nd Ed.)*, ISBN 0954161734, `http://www.gnu.org/software/gsl/`.
[9] F.A. Berends, H. Kuijf, B. Tausk, W.T. Giele, *On the Production of a W and Jets at Hadron Colliders*, Nucl. Phys. B357:32-64 (1991).
[10] R. Barlow, *Application of the Bootstrap resampling technique to Particle Physics experiments*, `http://www.hep.man.ac.uk/preprints/manhep99-4.ps` (2000).
[11] F. Maltoni, T. Stelzer, *MadEvent: Automatic Event Generation with MadGraph*, JHEP 302, 27 (2003).
[12] M.L. Mangano et al., *ALPGEN, a Generator for Hard Multiparton Processes in Hadronic Collisions*, JHEP 307, 1 (2003).
[13] M.G. Bowler, *e+ e- Production of Heavy Quarks in the String Model.*, Z. Phys C11, 169 (1981).
[14] C. Peterson *et. al*, *Scaling Violations in Inclusive e+ e- Annihilation Spectra.*, Phys. Rev. D27, 105 (1983).
[15] H. L. Lai *et al.* [CTEQ Collaboration], *Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions*, Eur. Phys. J. C 12, 375 (2000).
[16] J. Pumplin *et al.*, *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*, JHEP 0207, 012 (2002).